P$_{RW}$GEI: Poisson Random Walk based Gait Recognition

Pratheepan Yogarajah, Joan V. Condell and Girijesh Prasad
Intelligent System Research Centre (ISRC)
School of Computing and Intelligent Systems
University Of Ulster, Northern Ireland, United Kingdom

***Abstract*—Recently, gait recognition has received much increased attention from biometrics researchers. Most of the literature shows that existing appearance based gait feature representation methods, however, suffer from clothing and carrying object covariate factors. Some new gait feature representations are proposed to overcome the issue of clothing and carrying covariate factors, e.g. Gait Entropy Image (GEnI). Even though these methods provide a good recognition rate for clothing and carrying covariate gait sequences, there is still a possibility of obtain the better recognition rate by using better appearance based gait feature representations.**

To the best of our knowledge, a Poison Random Walk (PRW) approach has not been considered to overcome the issue of clothing and carrying covariate factors' effects in gait feature representations. In this paper, we propose a novel method, PRW based Gait Energy Image (P$_{RW}$GEI), to reduce the effect of covariate factors in gait feature representation. These P$_{RW}$GEI features are projected into a low dimensional space by a Linear Discriminant Analysis (LDA) method to improve the discriminative power of the extracted features. The experimental results on the CASIA gait database (dataset B) show that our proposed method achieved a better recognition rate than other methods in the literature for clothing and carrying covariate factors.

I. Introduction

Gait refers to the walking style of a human. Gaits have distinctive features that vary from one person to another. The way an individual normally walks is one of those distinctive features that may be used for recognition. Gait recognition methods can be mainly classified into three categories: spatiotemporal-based, model-based and appearance-based.

Spatiotemporal-based methods uncover gait shape variation information in both the spatial and temporal domains, and one such related work includes shape variation-based frieze features [1], [2]. Model-based methods [3], [4] aim to model the body and shape of the person when he/she is walking. Appearance-based methods focus on extracting the static (i.e. head and torso) and/or dynamic (i.e. motion of each arm, hand and leg) information of a walking person from sequences of binary silhouettes and representing the information as a single image. The computational cost of model-based methods is relatively high compared to appearance-based methods [5]. In the literature, cost effective appearance-based features are considered in gait recognition.

In appearance-based gait recognition, static features [6], dynamic features [7] and fusion of static and dynamic features [8] are used for gait recognition. From these works we can easily conclude that static and dynamic information of gait features are valuable information for gait recognition. Despite the fact that the above mentioned appearance-based features were shown by recent experiments to achieve promising gait recognition rates for normal (i.e. non-covariate) gait sequences, their use in the development of a biometric system is impractical. This is mainly because body-related parameters are not robust as they are dependent on clothing, bags, and other factors, [6], [9].

Figure 1 shows people appearing with normal clothing and different covariate factors. These covariate factors would certainly affect recognition performance.

Fig. 1. First row represents sample of individual walking sequences used for training and second row represents sample of individuals with different clothing and carrying objects used for testing, [21].

Sarkar et al. [11] described a baseline algorithm for gait recognition to examine the effects of the different covariates including viewpoint, footwear, walking surface, time, different clothing and different carrying objects. But Bouchrika et al. [12] argued that the work of [11] lacked exploratory analysis of the different gait features under covariate data due to the use of an appearance-based approach that includes covariate factors information.

On the other hand, in appearance-based methods, the Gait Energy Image (GEI) [10] has been proven to be the most effective. While GEI can attain good performance under normal gait sequences, it is sensitive to covariate factors, such as clothing and carrying of objects. To alleviate the effect caused by different clothing and carrying objects, Pratheepan et al. proposed covariate factor removal method [13], [14] on static features under the assumption that the static features can be extracted as clean silhouettes (i.e. without holes or noise). This assumption may not be applicable in all scenarios. Instead of removing covariate factors, different gait feature representations have been proposed in the literature [15], [16], [17], [18] to overcome the affect of different clothing and carrying object covariate factors, see Figure 2.

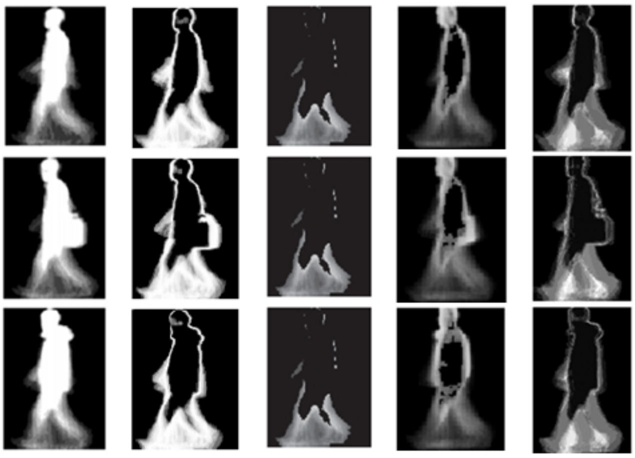

Fig. 2. shows different gait features for a particular individual from CASIA gait dataset. Columns from left to right represent GEI, GEnI, M_G, EGEI and AEI respectively. Rows from top to bottom represent normal, carrying bag and wearing coat conditions walking.

Yang et al. proposed an enhanced GEI (EGEI) [15] gait representation method to apply dynamic region analysis to improve dynamic information of the features extracted. But we can see that the major parts of bag and coat which should be removed are treated as static regions and remain in the EGEI, which indicates that they are treated as dynamic regions in EGEI.

Zhang et al. proposed an active energy image (AEI) [16]. The advantage of AEI is that it can retain the dynamic characteristics of gait for recognition. The active regions are the regions that would be calculated from the difference between two silhouette images in the gait sequence. AEI is created by summing up these active regions. In their experiments, they showed that an AEI has a higher recognition rate than GEI on the CASIA gait covariate dataset.

Shannon entropy based Gait Entropy Image (GEnI) [17] is used to reduce the covariate effect in gait features. Based on computing entropy, dynamic body areas which undergo a gait cycle will lead to high gait entropy value, whereas those areas that remain static would give rise to low values. On the other hand M_G [18] is generated by reducing as much static parts from GEI, see Figure 2. Experimental results of GEnI, M_G and AEI showed good recognition results for carrying and clothing covariate factors on the CASIA-B gait covariate dataset.

Even though these gait feature representations show a good recognition rate, their average recognition rates are not that promising (i.e. less than 75%). This indicates that a more robust appearance-based gait feature representation is needed to reduce the effect of clothing and carrying covariate factors and increase the recognition rate.

In this paper we propose a new gait feature representation to overcome the issue of carrying objects and different clothing covariate factors effects from silhouettes by using the Poisson Random Walk (PRW) approach. The overview of the approach is given in Section II. In Section III, details of feature extraction are described. Section IV provides details of the proposed recognition algorithm while Section V presents the experimental results. Finally Section VI gives the conclusion to the work.

II. Overview of the Approach

There are three main steps in our approach:

1) First extract the subject from each frame of the video sequence. The result is the silhouette of a person.
2) Secondly extract the PRW appearance-based features that reduce the effect of covariate factors. Then the LDA is applied to reduce the dimension of the data and also to improve the discriminative power of the extracted features.
3) In the last stage the k-nearest neighbour (*k*-NN) is applied to classify the data and make a decision, i.e. decide the identity of a person.

III. Feature Extraction

The silhouettes are extracted using a Gaussian model based background estimation method [19]. Then the bounding box of the silhouette image in each frame is computed. The silhouette image is extracted according to the size of the bounding box and the extracted image is resized to a fixed size (128*100 pixels). The purpose of resizing is to eliminate the scaling effect. The resized silhouette is then aligned centrally with respect to its horizontal centroid. After preprocessing the gait period is estimated in an individual's walking sequences.

A. Gait period estimation

Since our proposed gait feature templates depend on the gait period, we must estimate the number of frames in each walking cycle. A single walking cycle can be regarded as that period in which a person moves from the mid-stance (both legs are close together) position to a double support position (both legs are far apart), then the mid-stance position, followed by the double support position, and finally back to the mid-stance position.

The gait period can then be estimated by calculating the number of foreground pixels in the lower half of the silhouette image [11]. In mid-stance position, the silhouette image contains the smallest number of foreground pixels. In double support position, the silhouette contains the greatest number of foreground pixels. The gait period is calculated using the median of the distance between two consecutive minima. Then we applied Poisson Random Walk to reduce the effects of covariate factors in each silhouette image.

B. Poisson Random Walk

To the best of our knowledge, the Poison Random Walk (PRW) [20] has not been used for covariate factor removal and gait recognition. Here we use PRW [20] to reduce the covariate factors effects of binary silhouettes for gait recognition. Consider a shape as a given silhouette S in a grid plane (a binary image) and ∂S a simple closed curve as its boundary. The PRW approach assigns a value to every pixel of the silhouette [20]. This value is in fact the expected number of steps taken (starting from the pixel) to hit the boundary. $U(x, y)$ can be computed recursively as follows: At the boundary of S, i.e., $(x, y) \in S$, $U(x, y) = 0$. At every point (x, y) inside S, $U(x, y)$ is equal to the average value of its immediate four neighbours plus a constant (representing the amount of time required to get to an immediate neighbour), i.e.,

$$U(x,y) = \frac{1}{4}(U(x+1,y) + U(x-1,y) + U(x,y+1) + U(x,y-1)) + 1 \tag{1}$$

This constant is set to one time unit. Note that (1) is a discrete form approximation of the Poisson equation:

$$\Delta U(x,y) = -\frac{4}{h^2} \tag{2}$$

with $\Delta U = U_{xx} + U_{yy}$ denoting the Laplacian of U and $\frac{4}{h^2}$ denoting the overall scaling. For convenience, $\frac{4}{h^2}$ is set as 1 (intuitively, meaning one spatial unit per one time unit, where one spatial unit measures the distance to an immediate neighbour). Therefore, solve,

$$\Delta U(x,y) = -1 \tag{3}$$

with $(x, y) \in S$, subject to Dirichlet boundary conditions $U(x, y) = 0$ at the bounding contour ∂S. The pixels near the boundary have small values of U, see Figure 3. However, the gradient of U has larger values near the boundaries and smaller values in the center of the silhouette. We define the function Φ as follows:

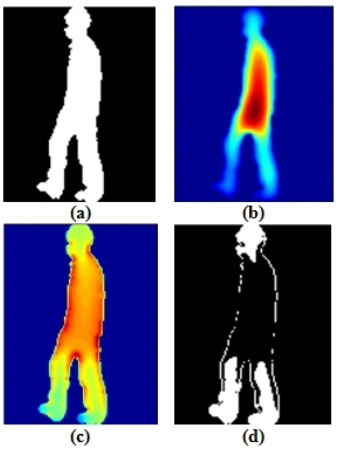

Fig. 3. (a) represents binary silhouette. (b) represents U of (a). (c) represents computed Ψ of (a). PRW$_{sil}$ binary silhouette is represented in (d).

$$\Phi(x,y) = U(x,y) + \parallel \nabla U(x,y) \parallel^2 \tag{4}$$

Φ has a distinctive characteristic to separate different parts of a shape based on their thickness. We consider $\Psi = \log(\Phi)$ to reduce the covariate factor effects. Then Ψ is scaled to make its values ranges from 0 to 255. Pixel coordinate positions corresponding to pixel values greater than 160 are selected. Then the selected pixels coordinate positions which correspond to the pixel values of Figure 3(a) are changed to 0 and the PRW silhouette (PRW$_{sil}$) is generated, see Figure 3(d).

Then a sequence of PRW$_{sil}$ for a gait period is considered to calculate the final P$_{RW}$GEI feature. P$_{RW}$GEI is calculated as follows:

$$P_{RW}GEI(x,y) = \sum_{n=1}^{N} PRW_{sil}^{n}(x,y) \tag{5}$$

where PRW_{sil}^{n} is a PRW_{sil} of the n^{th} frame of the particular gait cycle. N is the number of frames in a particular gait period. Figure 4 shows P$_{RW}$GEI features of sample of three individuals from CASIA-B database. Also Figure 5 shows feature representations of GEI, GEnI, M_G and our proposed P$_{RW}$GEI feature for a particular individual.

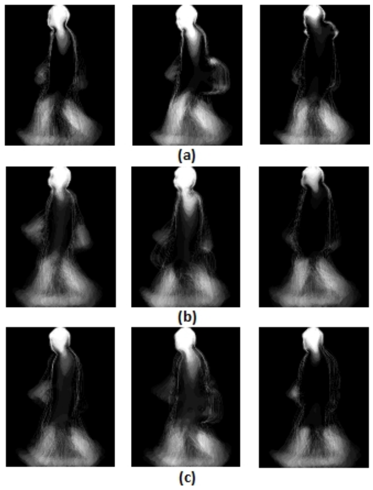

Fig. 4. (a), (b) and (c) show PRW-GEI for three individuals from CASIA gait database [21]. The columns from left to right represent P$_{RW}$GEI features for normal, different clothing and carrying objects covariate factors.

IV. Recognition

When gait sequences are represented as P$_{RW}$GEI, gait recognition can be performed by matching a *probe* P$_{RW}$GEI to the *gallery* P$_{RW}$GEI that has the minimal distance to the probe P$_{RW}$GEI. For ease of understanding, *gallery* refers to training data and *probe* refers to testing data. Suppose we have an N d-dimensional gallery P$_{RW}$GEI templates $\{x_1, x_2, ..., x_n, ..., x_N\}$ belonging to c different classes (i.e. individuals), where each template is a column vector obtained by concatenating the rows of the corresponding P$_{RW}$GEI which are subjected to Principal Component Analysis (PCA) [5].

PCA is an orthogonal linear transformation that transforms the data to a subspace of dimensionality $\tilde{d}$ (with $\tilde{d} < d$). The

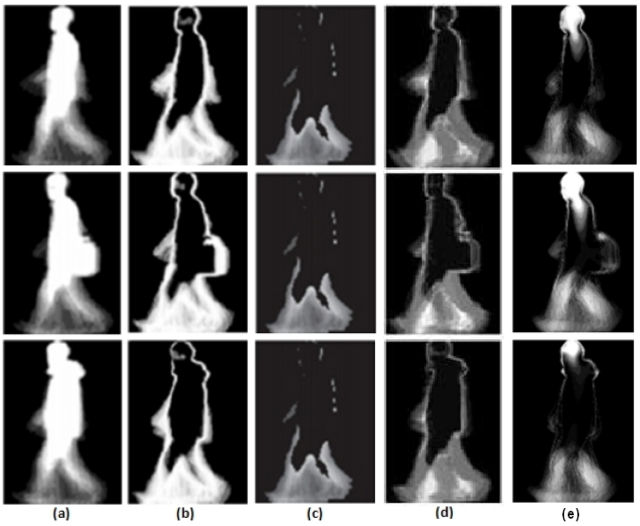

Fig. 5. (a), (b), (c), (d) and (e) show GEI, GEnI, M_G, AEI and our proposed feature respectively

PCA subspace keeps the greatest variances by any projection of the data so that the reconstruction error defined below is minimised:

$$J_{\tilde{d}} = \sum_{n=1}^{N} \left\| \left(m + \sum_{j=1}^{\tilde{d}} a_{nj} e_j \right) - x_n \right\|^2 \tag{6}$$

where m is the mean of the data, $\{e_1, e_2, ..., e_{\tilde{d}}\}$ are a set of orthogonal unit vectors representing the new coordinate system of the subspace, a_{nj} is the projection of the nth data to e_j. $J_{\tilde{d}}$ is minimised when $e_1, e_2, ..., e_{\tilde{d}}$ are the $\tilde{d}$ eigenvectors of the data covariance matrix with the largest eigenvalues (in decreasing order). Now the gallery template x_n is represented as a d-dimensional feature vector y_n and we have

$$y_n = [e_1, e_2, ..., e_{\tilde{d}}]^T x_n \tag{7}$$

PCA is followed by LDA which aims to find a subspace where data from different classes are best separated in a least square sense. Different from PCA, LDA is a supervised learning method which requires the gallery data to be labelled into classes. The LDA transformation matrix, W maximises

$$J(W) = \frac{|W^T S_B W|}{|W^T S_W W|} \tag{8}$$

where S_B is the between-class scatter matrix and S_W the within-class scatter matrix of the training (gallery) data in the PCA subspace $\{y_1, y_2, ..., y_n, ..., y_N\}$. $J(W)$ is maximised by setting the columns of W to the generalised eigenvectors that correspond to the $c-1$ nonzero eigenvalues in

$$S_B w_j = \lambda^j S_W w_j \tag{9}$$

where w_j is the jth column of W and c is the number of classes in the training data. Denoting these generalised eigenvectors as $\{v_1, v_2, ..., v_{c-1}\}$, a gallery template is represented in the LDA subspace as, [17]:

$$z_n = [v_1, ..., v_{c-1}]^T y_n \tag{10}$$

After this dimensionality reduction, both the gallery and probe P$_{RW}$GEI feature vectors are represented in a $(c-1)$ dimensional subspace and recognition can be computed as the distance of the probe feature vector to the gallery feature vector.

A. Classifier

The nearest neighbour classifier is adapted in the proposed algorithm. Suppose there are $N_{gallery}$ training subjects, for the individual recognition, N observations $\{\mathbf{x}_1, ..., \mathbf{x}_N\}$ can be represented as P$_{RW}$GEI images reshaped as column vectors. To perform recognition, we first obtain a set of "gallery" P$_{RW}$GEI for each class $\{\tilde{\mathbf{x}}_1, ..., \tilde{\mathbf{x}}_C\}$, where there are C classes, and find their projections onto the reduced subspace, $\{\tilde{\mathbf{z}}_1, ..., \tilde{\mathbf{z}}_C\}$. Then for a test observation we obtain the P$_{RW}$GEI $\tilde{\mathbf{x}}$, and calculate the distance between its projection onto the reduced subspace, $\tilde{\mathbf{z}}$, and each of the elements $\{\tilde{\mathbf{z}}_1, ..., \tilde{\mathbf{z}}_C\}$. The estimated class label i is the label of the gallery image which is the closest to the original image once projected to the LDA space:

$$i = \underset{j}{argmin} \parallel \tilde{\mathbf{z}} - \tilde{\mathbf{z}}_j \parallel \tag{11}$$

The similarity score represents the level of similarity between the testing data and the training data.

V. Results

We used the CASIA-B dataset [21] to evaluate the proposed algorithm. Dataset-B is a large covariate gait database. There are 124 subjects, and gait data was captured from 11 views. Three variations, namely view angle, clothing and carrying condition changes, are separately considered. We used the sequences collected at the 90^0 view (i.e. fronto parallel) with normal, clothing and carrying conditions for our experiments same as in [17], [16] and [18]. This is because the gait of a person is best brought out in the side view [22].

For each subject there are 10 gait sequences consisting of 6 normal gait sequences where the subject does not wear a bulky coat or carry a bag (CASIASetA), 2 carrying-bag sequences (CASIASetB) and 2 wearing-coat sequences (CASIASetC). The first 4 of the 6 normal gait sequences were used as the gallery set. The probe set included the rest of the normal gait sequences (CASIASetA2), CASIASetB and CASIASetC.

A. Results and Discussion with P$_{RW}$GEI

The performance of our P$_{RW}$GEI representation was compared with a direct template matching method [24], GEI , GEnI in [17], AEI in [16] and M_G in [18]. The result is given in Table I.

It can be seen from Table I that when the probe set (CASIASetA2) is tested with gallery set, all five methods yield good recognition rates. However, when the covariate conditions are different, the performance of all five methods degrade. Nevertheless, in the case of carrying objects gait sequences (CASIASetB) our P$_{RW}$GEI outperforms the rest. It shows a very impressive recognition rate, 93.1%. At the same time, in the case of wearing bulky coat gait sequences

TABLE I
PERFORMANCE ON CASIA-B (COVARIATE) DATASET

	[GEI+TM] [24]	[GEI+CDA] [17]	[GEnI+CDA] [17]	[AEI+PCA+LDA] [16]	[M_G^{ij}+CDA] [18]	[P$_{RW}$GEI+PCA+LDA] [proposed]
CasiaSetA2	97.6%	99.4%	98.3%	88.7%	**100**%	98.4%
CasiaSetB	52.0%	60.2%	80.1%	75.0%	78.3%	**93.1**%
CasiaSetC	32.7%	30.0%	33.5%	**57.3**%	44.0%	44.4%
Average	60.8%	63.2%	70.6%	73.7%	74.1%	**78.6**%

(CASIASetC), our P$_{RW}$GEI is comparable to the rest. The average recognition results, 78.6%, show that our P$_{RW}$GEI based gait recognition produced better recognition results than any other method shown in Table I.

It is noted that the selection of $\tilde{d}$ is effected by the dimensionality of LDA subspace, i.e. c-1. In particular, S_W becomes singular when $\tilde{d} < c$ or $\tilde{d} \gg c$. Therefore we did an experiment with $\tilde{d}$ ranging from 130 to 300 and $\tilde{d} = 140$ provided the better average recognition rate for normal, different clothing and carrying objects gait sequences, see Figure 6.

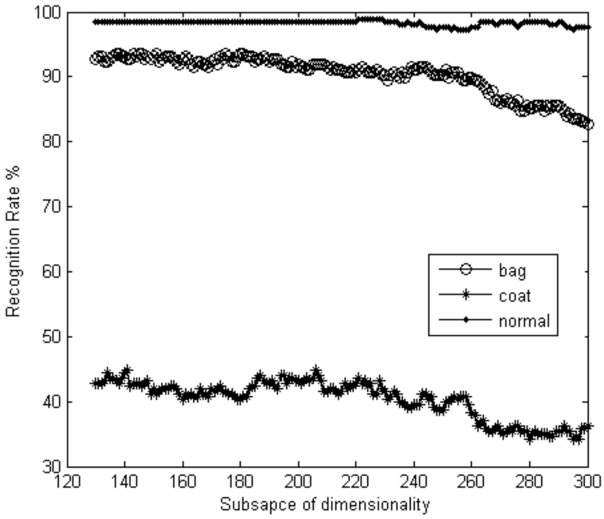

Fig. 6. Recognition rate on normal, clothing and carrying gait sequences against subspace of dimensionality - (PCA).

Increasing recognition rate by reducing the covariate effects in gait feature is the main concern of the above mentioned methods [17], [18], [16] and our method. Table II shows the performance comparison of our proposed gait feature representation and the other proposed methods. The dimensional reduction methods CDA and (PCA+LDA) are almost similar approaches. In our approach, we used 1-*KNN* for classification and it is a similar classification approach to above methods.

TABLE II
PERFORMANCE COMPARISON

Methods	Recognition Rate
Template Matching [24]	60.8%
GEI [17]	63.2%
GEnI [17]	70.6%
AEI [16]	73.7%
M_G [18]	74.1%
Proposed	**78.6**%

VI. CONCLUSION

In this paper, a novel P$_{RW}$GEI appearance based gait feature representation has been proposed for gait recognition. The performance of the proposed algorithm is evaluated experimentally using the CASIA-B dataset [21]. Experimental results showed that the performance of the proposed algorithm is promising for different clothing and comparable for carrying objects sequences.

Reducing covariate factor effects from gait sequences is the main concern in our method. Reducing covariate factor effects using our P$_{RW}$GEI based gait features worked very well. However we would like to investigate more robust methods of reducing the covariate factor effects as part of our future work. Apart from different clothing and carrying objects time and different views of gait sequences also have influences on gait recognition. Therefore time and view changes will be considered as future work.

Furthermore, LDA provides optimality for discrimination among the different classes. But some recent methods such as Geometric Mean for Subspace Selection [23], proved their effectiveness on class separability. Therefore as a future direction, we would like to investigate best class separable methods that can be used to optimise class separability and improve classification performance.

REFERENCES

[1] Y. Liu, R. Collins, Y. Tsin. (2001) Gait sequence analysis using frieze patterns, in: Proceedings of the Seventh European Conference on Computer Vision.

[2] S. Lee, Y. Liu and R. Collins. (2007) Shape variation-based frieze pattern for robust gait recognition, IEEE Conference on Computer Vision and Pattern Recognition , pp. 1-8.

[3] S.L. Dockstader, M.J. Berg and A.M. Tekalp. (2003) Stochastic kinematic modeling and feature extraction for gait analysis, IEEE Trans. Image Proc. 12(8), pp. 962-976.

[4] L. Wang, H. Ning, T. Tan and W. Hu. (2004) Fusion of static and dynamic body biometrics for gait recognition, IEEE Trans. Circuits and System

[5] L. Wang., T. N. Tan., W. M. Hu and H. Z. Ning. (2003), 'Automatic Gait Recognition Based on Statistical Shape Analysis', *IEEE Trans. on Image Processing*, Vol. 12(9), pp. 1120–1131.

[6] G. V. Veres, M. S. Nixon, and J. N. Carter. (2005) 'Modelling the Time-Variant Covariates for Gait Recognition'. *in Proc of 5th Intl conf on Audio-and Video-Based Biometric Person Authentication*, pp. 597–606.

[7] J. E. Cutting and D. Proffitt. (1981) 'Gait Perception as an Example of How we may Perceive Events'. *Intersensory perception and sensory integration*, pp. 249–273.

[8] T. Lam, R. Lee and D. Zhang. (2007) 'Human gait recognition by the fusion of motion and static spatio-temporal templates', *Pattern Recognition*, Vol. 40(9), pp. 2563–2573.

[9] D. K. Wagg and M. S. Nixon. (2004) 'On Automated Model-Based Extraction and Analysis of Gait'. *Conf Automatic Face and Gesture Recognition*, pp. 11–16.

[10] J. Han and B. Bhanu. (2006) 'Individual Recognition using Gait Energy Image'. *IEEE Transactions of PAMI*, Vol. 28(2), pp. 316–322.

[11] S. Sarkar., P.J. Phillips., Z. Liu., I.R. Vega., P. Grother., and K.W. Bowyer. (2005) 'The humanID gait challenge problem: data sets, performance, and analysis', *IEEE Trans. PAMI*, Vol. 27(2), pp. 162–177.

[12] I. Bouchrika., and M. Nixon. (2008) 'Exploratory Factor Analysis of Gait Recognition'. *In: 8th IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 1–6.

[13] Y. Pratheepan., J. V. Condell, and G. Prasad. (2009) 'The use of Dynamic and Static Characteristics of Gait for Individual Identification', *In Proceedings of the 13th International Machine Vision and Image Processing Conference (IMVIP), IEEE Computer Society*, pp. 111–116.

[14] Y. Pratheepan, Condell, J. V., Prasad, G. (2009) Individual Identification using Gait Sequences under different Covariate Factors. 7th International Conference on Computer Vision Systems (ICVS), pp 84-93.

[15] X. Yang, Y. Zhou, T. Zhang, G. Shu and J. Yang, (2008) Gait recognition based on dynamic region analysis, Signal Processing 88, pp. 2350-2356.

[16] E. Zhang, Y. Zhao and W. Xiong. (2010) Active energy image plus 2DLPP for gait recognition, Signal Processing 90 (7) , pp. 2295-2302.

[17] Bashir, K., Xiang, T., Gong, S. (2009) Gait recognition using gait entropy image. 3rd International Conference on Crime Detection and Prevention (ICDP), pp. 1–6.

[18] Khalid Bashir, Tao Xiang, Shaogang Gong. (2010) Gait recognition without subject cooperation. Pattern Recognition Letters 31(13): 2052–2060.

[19] C. Stauffer and W. E. L. Grimson. (1999) 'Adaptive background mixture models for real-time tracking', *in Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 2, pp. 246–252.

[20] L. Gorelick and M. Galun and E. Sharon and R. Basri and A. Brandt, 'Shape Representation And Classification Using The Poisson Equation', IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 28, No. 12, 2006.

[21] Casia gait database. (2010). http://www.cbsr.ia.ac.cn/english/Gait Databases.asp, online, September 2009.

[22] A. Kale, A. Sundaresan, A.N. Rajagopalan, N.P. Cuntoor, A.K. Roy Chowdhury, V. Kruger, and R. Chellappa. (2004) Identification of humans using gait. IEEE Transactions on Image Processing, 13(9):1163–1173.

[23] D. Tao, X. Li, X. Wu, S. J. Maybank. (2009) 'Geometric Mean for Subspace Selection', *IEEE Trans. Pattern Anal. Mach. Intell*, Vol. 31(2), pp. 260–274.

[24] Yu, S., Tan, D., Tan, T. (2006) A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition. In: ICPR, 441–444.